An Improved AES Hardware Trojan Benchmark to Validate Trojan Detection Schemes in an ASIC Design Flow

Sudeendra kumar K, Rakesh Chanamala, Sauvagya Ranjan Sahoo, K.K.Mahapatra
National Institute of Technology, Rourkela, India
(kumar.sudeendra, rakemtl, sauvagya.nitrkl, kmaha2)@gmail.com

***Abstract-* The semiconductor design industry has globalized and it is economical for the chip makers to get services from the different geographies in design, manufacturing and testing. Globalization raises the question of trust in an integrated circuit. It is for the every chip maker to ensure there is no malicious inclusion in the design, which is referred as Hardware Trojans. Malicious inclusion can occur by an in-house adversary design engineer, Intellectual Property (IP) core supplied from the third party vendor or at untrusted manufacturing foundry. Several researchers have proposed hardware Trojan detection schemes in the recent years. Trust-Hub provides Trojan benchmark circuits to verify the strength of the Trojan detection techniques. In this work, our focus is on Advanced Encryption Standard (AES) Trojan benchmarks, which is most vulnerable block cipher for Trojan attacks. All 21 Benchmarks available in Trust-hub are analyzed against standard coverage driven verification practices, synthesis, DFT insertion and ATPG simulations. The analysis reveals that 19 AES benchmarks are weak and Trojan inclusion can be detected using standard procedures used in ASIC design flow. Based on the weakness observed design modification is proposed to improve the quality of Trojan benchmarks. The strength of proposed Trojan benchmarks is better than existing circuits and their original features are also preserved after design modification.**

***Keywords:* Hardware Trojan, AES, Trust-Hub, Security, ASIC.**

I. INTRODUCTION

Due to time to market pressures and competition, semiconductor supply chain has changed significantly in the last decade. The service providers to these firms are located in different parts of the world. Because of globalization of the process, designs and devices are becoming vulnerable to malicious activities. The adversary can introduce hardware trojans at various steps in design process. Hardware trojans can be introduced into design at untrusted foundry during manufacturing. Intention of Trojan insertion is to destroy the system at future time or leak the confidential information like secret keys used in cryptographic engines. The insertion of trojans raised serious concerns regarding possible threats to defence systems. US Defence Science Board reports confirm the malicious insertions into the chips used in military systems [1]. Defence Departments of major economies in the world are paranoid about hardware trojans in the ICs used in their defence electronic equipments. [1][2]

The most vulnerable designs for hardware Trojan attack are cryptographic cores, processors and communication controllers transmitting sensitive data [2]. In cryptographic cores, Advanced Encryption Standard (AES) is widely used and most vulnerable cryptographic algorithms to hardware trojans. Side channel attacks are well known to hardware security community [4]. The recent entry into list is insertion of hardware trojans into AES designs at RTL coding stage and synthesized gate level netlist. The intention of malicious inclusions is to leak the key used in encryption. It is difficult to insert large trojans into design during manufacturing process [5]. It is relatively easy to insert Trojans at RTL design, in synthesized gate level netlist and in layout design. The study of variety of trojans, their taxonomy, affects, detection techniques, defence mechanisms are areas of interest for researchers. Good amount of literature is available on taxonomy, detection techniques and defence mechanisms [2] [2]. Trojan Detection techniques which are available in literature are based on verification, testing, timing fingerprints and side-channel analysis [6] [7].

Hardware Trojan detection capabilities of newly proposed/developed techniques should be validated on some Trojan benchmark circuits. Trust-hub [8] provides Hardware Trojan benchmarks for some standard processors, cryptographic cores, communication controllers etc. Strong and efficient benchmarks are necessary to validate the Trojan detection technique and to develop defence mechanisms against hardware trojans. Twenty one RTL benchmarks circuits for AES are available in Trust-hub. Side-channel analysis and timing fingerprints are mostly used to detect trojans on fabricated chips [7]. So it is always cost effective to detect trojans inserted at RTL level in pre-silicon stages like synthesis, functional verification, formal verification and DFT insertion which are part of standard ASIC design flow.

In this work, we use synthesis, functional verification and DFT insertion procedures followed in standard ASIC design flow to study the Trojan detection capabilities of these standard procedures. The aim of this work is also to study the strength and weakness of the Trust-hub AES RTL benchmark circuits. A novel Trojan is also presented which overcomes the weaknesses of the available Trust-hub benchmark circuits. AES is the most vulnerable cryptographic technique to hardware trojans and Trust-hub provides highest number of benchmarks for AES when compared to other designs. So AES is chosen to this study.

The rest of the paper is arranged as follows: Section II discusses the threat of hardware Trojan, pre-silicon and post silicon techniques available to detect hardware trojans. Section III describes AES benchmark circuits from Trust-hub, study and analysis of AES benchmarks. The strength and weaknesses of the AES benchmarks are discussed in section IV, section V describes the novel hardware Trojan designed to overcome the weaknesses of the existing benchmarks. Finally, section VI concludes this paper.

II. Hardware trojan and its detection

The structure of hardware Trojan includes two main parts, trigger and payload. Taxonomy of trojans and few detection schemes are discussed in [2].

Trojan detection approaches: Trojans get activated by rare signals and there are variety of trojans makes it tedious to devise a single silver bullet detection technique, which can find all trojans. At RTL level, a rogue designer can insert trojans in code and soft IP core in RTL obtained from third party vendor may have hardware trojans. The logic testing based Trojan detection techniques are proposed in [11] with an aim to activate the Trojan with rare inputs and comparing the results with golden values on Automated Test Equipment (ATE). According to [11], generating the test patterns to detect trojans is computationally infeasible. The Trojan detection approaches based on side-channel analysis is discussed in [11]. The parameters like path delays, leakage and transient currents are compared with data of golden chips which are fabricated in the trusted environment. It is difficult to find small trojans in deep submicron designs due to process variations and noise using side channel approach [11] in which the probability of false positive is high. Devising the appropriate logic testing and side-channel approach methods to find the small trojans inserted at manufacturing is still a difficult problem.

The focus of this work is to examine the response of Trust-hub benchmarks during synthesis, verification and DFT insertion. Before this endeavour, several researchers applied functional and formal verification techniques to ensure security against hardware trojans through defining security properties. These properties are applied on RTL code or on a third party IP core from a trusted team of engineers to ensure security [9][11]. Verification methods check RTL implementation for intended functionality. The challenge in hardware Trojan detection is that verification should check for any unintended behavior of the RTL implementation, which is relatively new to verification engineers. A case study on detecting the hardware trojans in third party IP cores is presented in [9]. The work in [9] will make use of code coverage, system Verilog assertions to discover the trojans in third party IP cores. [9] studies the trustworthiness of third party IPs using assertions, code coverage analysis, ATPG and equivalence checking methods and removes suspicious signals. [10] studies the 18 AES hardware Trojan benchmarks from Trust-hub and analyze area, leakage power and dynamic power and concludes that variation in area and power between Trojan inserted and Trojan free circuit is minute.

This work is an extension to [9] and [10] and our contribution is detailed analysis of strength and weakness of all the 21 AES benchmarks from Trust-hub. In [9], functional coverage, synthesis, gate level simulations, violations during DFT insertion is not used with an objective of Trojan detection. We make use of these techniques for Trojan detection. Our aim is to discover the strength and weakness of the AES benchmark circuits. AES is chosen to this study, because it is most vulnerable to hardware Trojan attacks. Similar study on AES Benchmark circuits from Trust-hub is presented in [10], but the work analyzes only area and power. Based on the weaknesses discovered in Trojan Benchmarks, we present novel Trojan inserted AES benchmarks overcome the weaknesses of existing benchmarks.

III. Trust-hub AES benchmark circuits

There are 21 AES Trojan Benchmark circuits from Trust-hub. There are three types of Trojan benchmarks designed for Denial of Service (DoS) and other 18 Trojan benchmarks promote side-channels to leak the confidential information like key. The 12 Trojans (AES-T100, AES-T200, AES-T300, AES-T700, AES-T800, AES-T900 AES-T1000, AES-T1100, AES-T1200, AES-T1300, AES-T1400, and AES-T1500) are described in [12] and reference for another 3 Trojans (AES-T400, AES-T1600, AES-T1700) benchmarks is available in [13]. There is no reference mentioned for other 6 Trojan (AES-T500, AES-T600, AES-T1800, AES-T1900, AES-T2000, and AES-T2100) benchmarks in Trust-hub repository.

The description of all 21Trojan AES benchmarks is shown in Table I. The area and dynamic power of the Golden design (Trojan free benchmark) and 21 different Trojan infected is shown in Table II. The highest difference in area between golden design and Trojan infected benchmark is reported for AES-T700 (4941.34 square microns). The change in area for few circuits is very minimal. In the era of deep submicron designs with multi-million gates, the reported difference in area in not very significant. Small change in design constraints will also impact the area numbers. Insertion of Trojan will have very less impact on area. Dynamic power consumption of Trojan infected circuit and golden circuit is also presented. The dynamic power is calculated on RTL implementation using a testbench developed in Verilog. Same test bench with no modification in input stimulus vectors is applied on golden design and all 21 AES Trojan bench marks to calculate dynamic power. The highest difference between golden and infected design is observed for AES-T1600 and value is 3.1732 mw. Based on area and power values, the difference between golden and benchmarks is not significant to decide whether design is infected by Trojan or not. Synopsys Design Compiler is used to synthesize the RTL design. The TSMC 65nm standard cell libraries used to synthesize the design. Synopsys Power Compiler and VCS are used for dynamic power calculations and functional verification of the circuits. Dynamic power is recorded by intentionally triggering the Trojan in benchmark circuits. There was no significant difference between the dynamic power values, when Trojan is on and Trojan is off. The average dynamic power values are shown in the Table II, after running the simulations for 10 times with and without triggering the Trojan circuit with different input stimulus. The average dynamic power values are shown in the Table II, after running the simulations for 10 times with and without triggering the Trojan circuit with different input stimulus.

TABLE I
DESCRIPTION OF 21 AES BENCHMARKS FROM TRUST-HUB

Benchmark	Effect	Side-channel	Description
AES-T100	leak info	power	The Trojan use pseudo-random number generator (PRNG) to create a CDMA code sequence. CDMA sequence is forwarded to a leakage circuit to set up a covert power side-channel. PRNG initialized to pre-defined value. (always on Trojan)
AES-T200	leak info	power	Same as AES-T100 but, PRNG initialized to pre-defined text. (always on Trojan).
AES-T300	leak info	power	Trojan leaks a byte of the round key for each round of the key schedule through a leakage circuit.
AES-T400	leak info	RF- signal	Modulating an unused pin on a chip generates an RF signal to transmit the key bits. Trojan gets activated with predefined input plaintext.
AES-T500	DoS	NA	Trojan is triggered by pre-defined input and drains battery within a short duration.
AES-T600	leak info	Leakage current	When a specific plaintext is given as input, the Trojan leaks the secret key through the leakage current. The leakage circuit consists of a shift register holding the secret key.
AES-T700	leak info	power	Same as AES-T100 but, Trojan trigger when pre-defined input is observed.
AES-T800	leak info	power	Same as AES-T100 but, Trojan trigger when pre-defined input is observed.
AES-T900	leak info	power	Same as AES-T100 but, Trojan trigger after 2^{128} encryptions.
AES-T1000	leak info	power	Same as AES-T100 but, Trojan trigger when pre-defined input is observed.
AES-T1100	leak info	power	Same as AES-T100 but, Trojan trigger when pre-defined input is observed.
AES-T1200	leak info	power	Same as AES-T100 but, Trojan trigger after 2^{128} encryptions.
AES-T1300	leak info	power	Trojan is triggered by specific input, it leaks key through increase in dynamic power.
AES-T1400	leak info	power	Trojan is triggered by input of specific sequence, it leaks key by increase in dynamic power.
AES-T1500	leak info	power	Trojan is triggered after 2^{128} encryptions, leaks key through increase in dynamic power.
AES-T1600	leak info	RF-Radio signal	Same as AES-T400 but, Trojan trigger when pre-defined input is observed.
AES-T1700	leak info	RF-Radio signal	Same as AES-T400 but, Trojan trigger after 2^{128} encryptions
AES-T1800	DoS	NA	Trojan is triggered by pre-defined input and drains battery within a short duration.
AES-T1900	DoS	NA	Trojan is triggered after 2^{128} encryptions and drains battery within a short duration.
AES-T2000	leak info	leakage current	Trojan is triggered by pre-defined input and increases leakage current.
AES-T2100	leak info	leakage current	Trojan is triggered after pre-defined number of encryptions and increases leakage current.

DoS: Denial of Service; NA: Not Applicable;

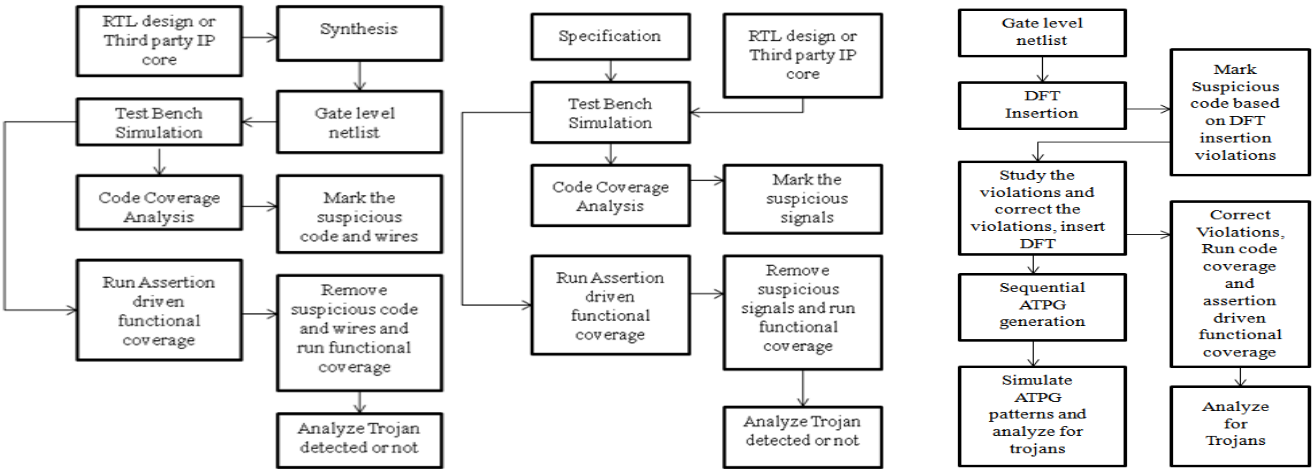
RTL design or Third party IP core
Synthesis
Test Bench Simulation
Gate level netlist
Code Coverage Analysis
Mark the suspicious code and wires
Run Assertion driven functional coverage
Remove suspicious code and wires and run functional coverage
Analyze Trojan detected or not
Specification
RTL design or Third party IP core
Test Bench Simulation
Code Coverage Analysis
Mark the suspicious signals
Run Assertion driven functional coverage
Remove suspicious signals and run functional coverage
Analyze Trojan detected or not
Gate level netlist
DFT Insertion
Mark Suspicious code based on DFT insertion violations
Study the violations and correct the violations, insert DFT
Correct Violations, Run code coverage and assertion driven functional coverage
Sequential ATPG generation
Simulate ATPG patterns and analyze for trojans
Analyze for Trojans

Fig. 1. (a) Verification flow at RTL level (b) Gate-level simulations and verification flow (c) DFT Insertion and ATPG simulation

TABLE II
AREA AND DYNAMIC POWER OF 21 AES BENCHMARKS

GOLDEN (TROJAN FREE) BENCHMARK AREA: 379280.9 SQUARE MICRONS
GOLDEN (TROJAN FREE) BENCHMARK DYN.POWER:184.5204 MILLIWATT

Benchmark	Dynamic Power	Power Delta	Area	Area Delta
AES-T100	185.0515	0.5311	379917.728	636.828
AES-T200	185.1306	0.6102	379936.448	655.548
AES-T300	184.5204	0	379280.888	-0.01199
AES-T400	184.7712	0.2508	380094.488	813.588
AES-T500	184.5204	0	379280.888	-0.01199
AES-T600	184.5315	0.0111	379408.328	127.428
AES-T700	187.481	2.9606	384222.2482	**4941.348**
AES-T800	187.3451	2.8247	384113.1681	4832.268
AES-T900	185.004	0.4836	381290.048	2009.148
AES-T1000	187.3202	2.7998	383795.2882	4514.388
AES-T1100	184.9933	0.4729	384096.9682	4816.068
AES-T1200	184.9933	0.4729	381311.648	2030.748
AES-T1300	184.5068	-0.0136	379397.168	116.268
AES-T1400	186.8748	2.3544	383387.0482	4106.148
AES-T1500	184.5225	0.0021	380389.328	1108.428
AES-T1600	187.6936	**3.1732**	384101.1282	4820.228
AES-T1700	184.6342	0.1138	382709.528	3428.628
AES-T1800	186.8379	2.3175	379280.888	-0.01199
AES-T1900	184.5204	0	379280.888	-0.01199
AES-T2000	184.5574	0.037	379694.888	413.988
AES-T2100	184.5159	-0.0045	380477.168	1196.268

Power in milliwatt; Area in square microns; Delta: numerical difference between golden and trojan infected benchmark circuit.

It is difficult to judge whether the third party IP core or RTL code is infected by Trojan based on the dynamic power consumption.

IV. STRENGTH AND WEAKNESS OF AES BENCHMARKS

The figure 1 shows three different flows followed in this work. All the three flows are followed in general ASIC design [20]. The Benchmarks can be broadly categorized as follows: always on trojans and conditionally triggered trojans. There are three Trojan benchmarks (AES-T100, AES-T200, and AES-T300) which are always on and rest 18 benchmarks are conditionally triggered. Based on the intention of insertion, benchmarks are classified as: Denial of Service (DoS) and information leaking trojans.

A.Denial of Service(DoS):

There are three Trojan benchmarks (AES-T500, AES-T1800 and AES-T1900) in this category. Intention of these trojans is to drain battery at faster rate once triggered. All these trojans are conditionally triggered by an event or by specific input. These trojans are dangerous for lightweight AES implementations generally used in battery operated medical devices like pacemakers etc. Trojan consists of shift register which rotates the data continuously, to drain the battery. These trojans are strong enough and go undetected in RTL verification (figure 1(a)). A blank Verilog module with no inputs, outputs and gates gets created in gate level netlist generated out of synthesis (means some block in the design is not synthesized). Careful observation of the schematic view of gate level netlist in synthesis tool and comparing with RTL code, it points to the shift register block inside the benchmark circuit, which is used in draining the battery. The code coverage analysis during gate level simulations will point out this blank Verilog module as un-triggered. The blank Verilog module looks suspicious and gate level simulations will pass the functional verification with 100% coverage after removal of blank Verilog module. This weakness is observed in all three DoS benchmarks.

The information leaking benchmarks can be categorized based on the covert side-channel created by hardware Trojan to leak secret keys. The classification is as follows: using CDMA based covert power side-channel to leak information, increasing the leakage current to leak information, leaking information by increasing the dynamic power consumption and leaking key by transmitting RF signal on an unused pin (acts like antenna), such that AES keys can be received on AM Radio easily.

B.Leakage current based trojans:

There are six Trojan benchmarks (leakage current: AES-T600, AES-T2000 and AES-T2100) in this category. Trojan leaks the secret key through leakage current. The leakage circuit consists of a shift register and two inverters. Shift register hold the secret key. The LSB is connected to one inverter which is an input to the other inverter. When LSB of shift register is ‘0’, a path between power and ground is created by the PMOS of the first inverter and NMOS of the second inverter for limited time. This increases the leakage current. The adversary can easily determine the bits of key by measuring leakage current. The leakage circuit is common to all three benchmarks. The shift register and two inverters block is designed with only inputs and no outputs like DoS benchmarks. These benchmarks will also create blank Verilog module in gate level netlist (leakage circuit not synthesized). These three trojans get detected similarly like DoS benchmarks during gate-level simulations.

C. RF Signal based trojans:

This category has got three (AES-T400, AES-T1600 and AES-T1700) trojans. An unused pin on a chip is used to generate the RF signal. Signal is used to leak the secret key. The data can be received at 1560 KHz with an AM radio. A single beep followed by a pause represents a ‘0’ and double beep followed by a pause represents ‘1’. It is very rare in a design to get unused pins. The synthesis tool will generate warnings about unused pin. The code coverage (toggle coverage) will not covered completely for all pins until Trojan gets triggered. Once Trojan gets triggered it can be easily gets detected in RTL verification. The RTL verification flow (toggle coverage with functional coverage) will detect this category of trojans.

D. CDMA based power side-channel based trojans:

There are six trojans (AES-T700, AES-T800, AES-T900, AES-T1000, and AES-T1200) in this category. When a predefined input is observed, the trojans leaks the secret key through a covert channel. Covert channel is will make use of CDMA to distribute the leakage of single bits over several clock cycles. Pseudo Random Number Generator (PRNG) is used to create CDMA sequence. CDMA sequence is used for XOR modulation of secret bits. The modulated sequence is forwarded to leakage circuit, which forms a covert power side-channel. Leakage circuit is designed by connecting eight flip-flops and XOR gate at the end with large capacitance.

This category of trojans is strong enough to bypass RTL verification and post synthesis gate level simulations. During DFT insertion, DFT violations and errors generated from the tool gives hint on asynchronous blocks, which are suspicious. After removing the suspicious asynchronous blocks, all benchmarks will pass the functional coverage. Converting all asynchronous blocks into synchronous blocks will solve the problem and scan chains are inserted. ATPG tool will generate the patterns (stuck-at-fault and path delay fault) which trigger trojans. Secret key leakage is observed during ATPG simulations. With this proof, trojans get detected in this set of benchmarks.

Trojan trigger block is designed in an asynchronous way. The sensitivity list of always block inside the trojan trigger circuit makes the trojan asynchronous and DFT insertion errors will get generated in DFT insertion tools.

E.Dynamic power side-channel trojans:

There are 3 trojans (AES-T1300, AES-T1400 and AES-T1500) in this category. Trojan gets triggered when a pre-defined input is observed. Trojan is created by artificially introduce leaking intermediate states in key schedule that depend on known input bits and key bits, which is not occur during normal processing of cipher. Trojan leaks the AES key for each key schedule. The leakage circuit is shift register loaded with an alternating sequence of zeros and ones. Shift register is enabled when input to leakage circuit is one, which results in additional dynamic power consumption. Adversary can measure the power and decode the key from the power analysis.

The weakness of these benchmarks is similar to DoS and leakage current. Shift register leakage circuit common across these three types and creates blank Verilog module which get detected in coverage driven gate level simulations.

F.Always on trojans:

The three trojans AES-T100, AES-T200 and AES-T300 are always on trojans. AES-T300 will easily get detected during gate level simulations like DoS, leakage current and dynamic power trojans. The trojans AES-T100 and AES-T200 are completely synchronous in design and hardware Trojan will go undetected in all three flows described in figure 1. The first two trojans are strong AES Trojan benchmarks. Side channel based Trojan detection methods [11] is required detect trojans in AES-T100 and AES-T200. In this work, we have not focused on side-channel methods of Trojan detection.

The 12 Trojans (AES-T100, AES-T200, AES-T300, AES-T700, AES-T800, AES-T900 AES-T1000, AES-T1100, AES-T1200, AES-T1300, AES-T1400, AES-T1500) are described in [18] are designed for FPGA. Except AES-T100 and AES-T200, other benchmarks are weak and trojans are detected in straight forward verification techniques and DFT insertion procedures. From the above discussion, it is evident that, researchers need strong benchmarks to verify and validate novel hardware Trojan detection techniques for ASIC design.

V. Novel Hardware Trojan

Except AES-T100 and AES-T200 benchmarks, other benchmarks will get detected by standard verification and DFT procedures. The major weakness in the AES benchmarks is shift register design for DoS, leakage current and dynamic power side-channel benchmark circuits. The weakness of the shift register is there is no output port. This creates blank Verilog module (un-synthesized), which gets detected in verification flows. Tweaking the design by adding the output port to the shift register part in the benchmarks by without disturbing the features of the benchmarks, Trojan benchmarks will become stronger. The presence or removal of shift register (identified as TSC in most of the Trojan benchmarks) has no influence on functionality of AES. In all benchmarks at shift register (TSC), easily bypasses code coverage and functional coverage techniques and get caught in synthesis. The novel hardware Trojan is proposed by adding the output port to the shift register without affecting the functionality of AES and not disturbing the originality of a benchmark (promoting various types of side-channels like leakage current, increase in dynamic power consumption).

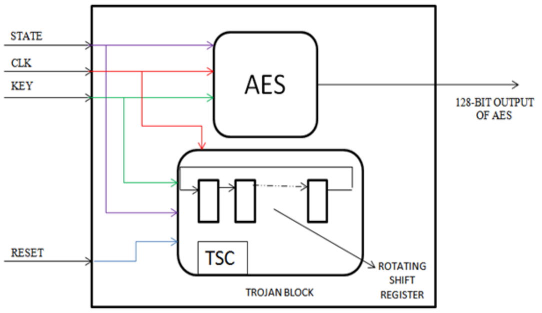

Fig 2 Structure of AES Trojan Benchmark of DoS, Leakage current and Dynamic power side-channel Trojan

The general structure of the AES Trojan Benchmarks (DoS, Leakage current and dynamic power side channel type is shown in figure 2. From the figure it is evident that, Trojan block has got no outputs and synthesis tools ignore Trojan block and no netlist will get synthesized. The existing benchmarks with similar structure are DoS, Leakage current and dynamic power side channel type. The above benchmarks can be modified as shown in figure 3. The Benchmarks are modified by adding an output port. A new design with output port in Trojan block is created with no affect on the AES functionality and also the original characteristics of benchmarks are also not affected. After modification, design is verified. The comparison of the RTL simulation of the modified benchmark (all 3 types: DoS, Leakage current and Dynamic power side channel) and original benchmark shows no difference.

The modified benchmark circuits are also verified using Xilinx Spartan 3E FPGA board and results are compared with original benchmarks. Denial of Service (DoS) and Dynamic power side channel benchmarks are verified on FPGA. The characteristics of modified benchmark circuits are similar to original benchmarks. Leakage current benchmarks are not verified on FPGA. The shift register intermediate signals are taken out as 128-bit signal, fed into multiplexer input and select line of mux shown in figure 3 is also 128 bit width. The state and key is directly taken into Trojan block and shift register rotates the state and key values after performing few combinational operations like XOR etc. This will leak key in the form of dynamic power with the increase in the switching activity. The Trojan is triggered by predefined input or condition. In the select line of multiplexer, all 128 bits should become zero to push the output of shift register as encrypted output. By the AES structure and operation, the probability of 128-bits getting zero value is very less in the shift register. So the result of AES block will get connected to output.

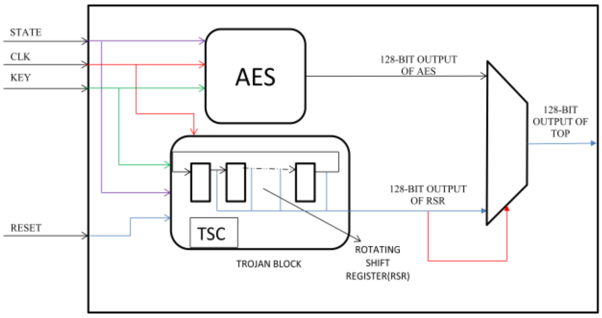

Fig 3 Structure of modified or New AES Trojan Benchmark of DoS, Leakage current and dynamic power side-channel Trojan

The code coverage driven simulations were not able to detect trojans in benchmarks after modification. After synthesis, Trojan circuit gets synthesized and code coverage driven gate level simulations are not able to detect the trojans in synthesized netlist. 100% coverage (exhaustive coverage) is achieved in line coverage, statement coverage and branch coverage using different stimulus sets in the testbench. Trojan will go undetected in line, statement and branch coverage. The exhaustive toggle coverage of internal signals of the design detected the Trojan, by directly affecting the functionality of the AES. To achieve toggle coverage 100%, we have to make all select lines of mux to zero. In most of the cases, the general practice in verification methodologies is to get 100% functional coverage and getting 95% and above of different code coverage (line, branch, and toggle) is fair. The novel Trojan will go undetected during DFT insertion without raising any warnings or errors. So the probability of Trojan getting detected is very less in general verification methodologies. By incorporating the above modification, the Trust-hub benchmarks AES-T500, AES-T1800 and AES-T1900 (DoS), AES-T600, AES-T2000 and AES-T2100 (Leakage current) and AES-T1300, AES-T1400 and AES-T1500 (Dynamic power side channel Trojans) will become stronger circuits against general verification methodologies and DFT insertion procedures.

VI. Conclusions

In this paper, all 21 AES Trojan benchmarks from Trust-Hub are studied exhaustively and their strengths and weaknesses are discussed. It is proved that, it is difficult to use the area and dynamic power consumption parameters to predict or detect the Trojans in the circuit. Further, coverage driven verification techniques used in general practice, Synthesis, DFT insertion and ATPG simulations using industry standard tools are applied on Trojan benchmarks to check the possibility of Trojan detection. Except AES-T100 and AES –T-200, in all other 19 benchmarks Trojan will get detected. Based on the weakness observed in the benchmark, improvement in the design to mitigate the weakness is proposed to make the benchmarks strong. The proposed design modification to existing weak benchmarks is proposed without disturbing the original features. The proposed novel hardware trojan benchmark is verified.

References

[1] "Report of the Defense Science Board Task Force on High Performance Microchip Supply", Defense Science Board, US DoD, Feb.2005; http://www.acq.osd.mil/dsb/reports/2005-02-HPMS_Report_Final.pdf.

[2] M.Tehranipoor, Farinaz Koushanfar, A survey of Hardware Trojan Taxonomy and Detection, IEEE Design and Test of Computers-2010.

[3] Stefan Mangard, Elisabeth Oswald, Thomas Popp, "Power Analysis attacks- revealing secrets of smart cards", Springer Science-2007.

[4] Y.Jin and Y.Markis "Hardware trojan detection using path delay fingerprint", IEEE intl. workshop on Hardware oriented security and trust, pp.51-57,2008.

[5] M.Bilzor et.al, "Evaluating Security requirements in a general purpose processor by combining assertion checkers with code coverage" IEEE International Symposium on HOST, San Francisco, CA, June 2012, Pages 49-54.

[6] Seetharaman Narasimahan, *et.al,* Hardware Trojan Detection by Multiple Parameter Side-Channel Analysis, IEEE Transactions on computers, vol.62, No.11, November-2013.

[7] https://www.trust-hub.org/resources

[8] Xuehui Zhang and Mohammad Tehranipoor, Case Study*:* Detecting Hardware Trojans in Third-Party Digital IP Cores, IEEE International Symposium on HOST-2011.

[9] Trey Reece, William H Robinson, "Analysis of data-leak hardware Trojans in AES cryptographic circuits" Technologies for Homeland Security (HST), 2013 IEEE International Conference on pp. 467-472

[10] L. Lin, M. Kasper, T. Güneysu, C. Paar and W. Burleson, "Trojan Side-Channels: Lightweight Hardware Trojans through Side- Channel Engineering," 11th International Workshop Cryptographic Hardware and Embedded Systems (CHES), pp. 382-395, 2009.

[11] Alex Baumgarten, Michael Steffen, Matthew Clausman, Joseph Zambreno, "A case study in hardware Trojan design and implementation," International Journal of Information Security, Volume 10, Issue 1, pp. 1-14, 2011.